Available online at www.sciencedirect.com
SCIENCE DIRECT®
Economics Letters 81 (2003) 321–326
economics letters
www.elsevier.com/locate/econbase

Simplified marginal effects in discrete choice models

Soren Anderson[a], Richard G. Newell[b,*]

[a] University of Michigan, Ann Arbor, MI, USA
[b] Resources for the Future, 1616 P St. NW, Washington, DC 20036, USA

Received 27 February 2003; received in revised form 10 June 2003; accepted 24 June 2003

Abstract

We show that with a simple normalization of explanatory variables, marginal effects in probit and logit models simplify dramatically, becoming a function of only the estimated constant term. Related simplifications hold for computation of asymptotic variances of these effects.

Keywords: Discrete choice; Marginal effect; Normalization

JEL classification: C25; C81

1. Introduction

It is well known that parameter estimates from discrete choice models, such as probit and logit, must be transformed to yield estimates of the marginal effects—that is, the change in predicted probability associated with changes in the explanatory variables (see, for example, Greene, 2003, p. 667). The marginal effects are nonlinear functions of the parameter estimates and the levels of the explanatory variables, so they cannot generally be inferred directly from the parameter estimates. Some statistical packages will not directly compute these effects, forcing users to program these procedures themselves. Beyond any computational issues, we believe the approach we suggest builds intuition and clarifies the relationship between discrete choice parameter estimates and their associated marginal and dummy variable effects.

We show that after a simple normalization of the explanatory variables so that they equal 0 at the desired reference point, marginal effects for continuous variables simplify dramatically, becoming a function of only the estimated constant term. We present similar simplifications for computation of the

* Corresponding author. Tel.: +1-202-328-5111; fax: +1-202-939-3460.
E-mail address: newell@rff.org (R.G. Newell).

0165-1765/$ - see front matter
doi:10.1016/S0165-1765(03)00212-X

asymptotic variance of marginal effects, as well as for the effects of dummy variables on predicted probabilities.

2. Normalization of explanatory variables

The first step in this simplification is to center all continuous variables at the desired reference point. The most commonly chosen reference point for calculating marginal effects in models with non-linear explanatory variables is at the variable means. Taking deviations from means will yield a 0 value for the normalized variable at the mean of the original variable. If a log form is being used, the variable can first be normalized to equal one at the mean (i.e., by dividing by the variable's mean), and then logs can be taken so that the variable equals 0 at its logged mean.[1] Of course, the reference point is not limited to variables means; variables can be normalized to equal 0 at any desired value. Categorical or dummy variables can be treated similarly, and should be coded (i.e., as 0 or 1) so that the constant term corresponds to the desired reference group for which the marginal effects will be calculated (i.e., the omitted group).[2]

3. Marginal and discrete effects of explanatory variables

The predicted probability from a binary choice model is given by

$$E[y \mid \mathbf{x}] = F(\boldsymbol{\beta}'\mathbf{x}), \tag{1}$$

where y is a choice variable, $\mathbf{x}$ is a vector of explanatory variables, $\boldsymbol{\beta}$ is a vector of parameter estimates, and F is an assumed cumulative distribution function. Assuming F is the standard normal distribution (Φ) produces the probit model, while assuming F is the logistic distribution (Λ) produces the logit model, where $\Lambda\,(\boldsymbol{\beta}'\mathbf{x}) = \exp(\boldsymbol{\beta}'\mathbf{x})/(1+\exp(\boldsymbol{\beta}'\mathbf{x}))$.

In these models, marginal effects for continuous variables (i.e., the marginal changes in expected probability $\partial E[y]/\partial \mathbf{x}$) are equal to

$$\partial E[y \mid \mathbf{x}]/\partial \mathbf{x} = f(\boldsymbol{\beta}'\mathbf{x})\boldsymbol{\beta}, \tag{2}$$

where f is the corresponding probability density function. For the probit model f is given by ϕ, the standard normal density function, where

$$\phi(\boldsymbol{\beta}'\mathbf{x}) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(\boldsymbol{\beta}'\mathbf{x})^2\right), \tag{3}$$

[1] Under this model specification, marginal effects are interpreted as the change in predicted probability associated with *percent* changes in the continuous independent variables.

[2] Note that while model coefficients are invariant to centering of first-order terms, they are not invariant to centering of higher-order terms such as variable interactions or quadratic terms.

Table 1
Lookup table for computing marginal/discrete effects and their variances

Constant term, c	Distribution function $F(c)$ (predicted probability)		Density function $f(c)$ (marginal effect scale factor)	
	Logit $\Lambda(c)$	Probit $\Phi(c)$	Logit $\gamma(c)$	Probit $\phi(c)$
0.000	0.500	0.500	0.250	0.399
0.213	0.553	0.584	0.247	0.390
0.312	0.577	0.622	0.244	0.380
0.388	0.596	0.651	0.241	0.370
0.453	0.611	0.675	0.238	0.360
0.512	0.625	0.696	0.234	0.350
0.565	0.638	0.714	0.231	0.340
0.616	0.649	0.731	0.228	0.330
0.664	0.660	0.747	0.224	0.320
0.710	0.670	0.761	0.221	0.310
0.755	0.680	0.775	0.218	0.300
0.799	0.690	0.788	0.214	0.290
0.841	0.699	0.800	0.210	0.280
0.884	0.708	0.812	0.207	0.270
0.925	0.716	0.823	0.203	0.260
0.967	0.724	0.833	0.200	0.250
1.008	0.733	0.843	0.196	0.240
1.050	0.741	0.853	0.192	0.230
1.091	0.749	0.862	0.188	0.220
1.133	0.756	0.871	0.184	0.210
1.175	0.764	0.880	0.180	0.200
1.218	0.772	0.888	0.176	0.190
1.262	0.779	0.896	0.172	0.180
1.306	0.787	0.904	0.168	0.170
1.352	0.794	0.912	0.163	0.160
1.399	0.802	0.919	0.159	0.150
1.447	0.810	0.926	0.154	0.140
1.498	0.817	0.933	0.149	0.130
1.550	0.825	0.939	0.144	0.120
1.605	0.833	0.946	0.139	0.110
1.664	0.841	0.952	0.134	0.100
1.726	0.849	0.958	0.128	0.090
1.793	0.857	0.963	0.122	0.080
1.866	0.866	0.969	0.116	0.070
1.938	0.874	0.974	0.110	0.061
2.063	0.887	0.980	0.100	0.047
2.197	0.900	0.986	0.090	0.036
2.342	0.912	0.990	0.080	0.026
2.502	0.924	0.994	0.070	0.017
2.681	0.936	0.996	0.060	0.011
2.887	0.947	0.998	0.050	0.006
3.134	0.958	0.999	0.040	0.003
3.444	0.969	1.000	0.030	0.000
3.871	0.980	1.000	0.020	0.000
4.586	0.990	1.000	0.010	0.000
$\geq$ 7.621	1.000	1.000	0.000	0.000

For negative values of c, note that $f(-c)=f(c)$ and $F(-c)=1-F(c)$.

while for the logit model the logistic density function is given by

$$\gamma(\boldsymbol{\beta}'\mathbf{x}) = \Lambda(\boldsymbol{\beta}'\mathbf{x})(1 - \Lambda(\boldsymbol{\beta}'\mathbf{x})).$$

The density function $f(\boldsymbol{\beta}'\mathbf{x})$ can therefore be thought of as a scale factor that translates raw parameter estimates into marginal effects. The point of this note is to find simple forms for this scale factor, as well as analogous factors for the effects of dummy variables and the variances of these effects.

With all continuous variables normalized to equal 0 at the desired reference point, $\boldsymbol{\beta}'\mathbf{x}$ simplifies to c, and $f(\boldsymbol{\beta}'\mathbf{x})$ simplifies to $f(c)$, where c is the estimated constant term. As c gets increasingly positive, $F(c)$ approaches 1, $f(c)$ approaches 0, and the marginal effects therefore approach 0. Similarly, as c becomes increasingly negative, $F(c)$ approaches 0, and $f(c)$ and the marginal effects again approach 0. This pattern is shown in more detail in Table 1, which gives the full range of values of $f(c)$ for associated values of c. To convert parameter vector $\boldsymbol{\beta}$ to its associated marginal effects, one can simply multiply by the value of $f(c)$ for the estimated value of c. Note that because both distributions are symmetric, $f(-c)=f(c)$, and the predicted probability at $-c$ is given by $1-F(c)$.

The discrete effect of a dummy variable is found by taking the difference in the predicted probability with and without that dummy variable equal to 1. Given the normalizations described above, this results in the following simple relationship for the discrete probability effect of a dummy variable:

$$E[y \mid d=1] - E[y \mid d=0] = F(c+d) - F(c), \tag{5}$$

where d is the estimated parameter for the dummy variable. As c becomes increasingly positive, both the first and second terms of this expression approach 1, so the net effect of the dummy variable approaches 0. As c becomes increasingly negative, both the first and second terms approach 0 and, again, the net effect of the dummy variable approaches 0. The effect of a dummy variable for any value of c and d can be readily found using Table 1.[3]

4. Variances of marginal and discrete effects

Variances for marginal effects can be calculated using the linear approximation approach (delta method). The asymptotic covariance matrix for the marginal effects is given by

$$\text{Asy. Var}[\phi(\boldsymbol{\beta}'\mathbf{x})\boldsymbol{\beta}] = \phi(\boldsymbol{\beta}'\mathbf{x})^2[\mathbf{I} - (\boldsymbol{\beta}'\mathbf{x})\boldsymbol{\beta}\mathbf{x}']\mathbf{V}[\mathbf{I} - (\boldsymbol{\beta}'\mathbf{x})\boldsymbol{\beta}\mathbf{x}'] \tag{6}$$

for the probit model, and

$$\text{Asy. Var}[\gamma(\boldsymbol{\beta}'\mathbf{x})\boldsymbol{\beta}] = \gamma(\boldsymbol{\beta}'\mathbf{x})^2[\mathbf{I} + (1 - 2\Lambda(\boldsymbol{\beta}'\mathbf{x}))\boldsymbol{\beta}\mathbf{x}']\mathbf{V}[\mathbf{I} + (1 - 2\Lambda(\boldsymbol{\beta}'\mathbf{x}))\mathbf{x}\boldsymbol{\beta}'] \tag{7}$$

[3] Simply substitute $c+d$ for c to find the value of $F(c+d)$ from Table 1.

for the logit model, where $\mathbf{V}$ is the estimated asymptotic covariance matrix of $\boldsymbol{\beta}$ (Greene, 2003, p. 675). Given the above normalizations, one can show that the asymptotic variance for the marginal effect of a particular continuous coefficient estimate β is given by

$$\text{Asy. Var}[\phi(\boldsymbol{\beta}'\mathbf{x})\beta] = \phi(c)^2(\sigma_\beta^2 + c^2\beta^2\sigma_c^2 - 2c\beta\sigma_{\beta c}) \tag{8}$$

for the probit model, and

$$\text{Asy. Var}[\gamma(\boldsymbol{\beta}'\mathbf{x})\beta] = \gamma(c)^2(\sigma_\beta^2 + (1 - 2\Lambda(c))^2\beta^2\sigma_c^2 + 2(1 - 2\Lambda(c))\beta\sigma_{\beta c}) \tag{9}$$

for the logit model, where σ_c^2 is the asymptotic variance for the constant, σ_β^2 is the asymptotic variance for the estimate of parameter β, and $\sigma_{\beta c}$ is the asymptotic covariance between β and c.[4] One can now compute the asymptotic variance for the marginal effect using only Table 1 along with the estimated values of β, c, σ_β^2, σ_c^2, and $\sigma_{\beta c}$.

To calculate the asymptotic variance for the effect of a dummy variable, we begin with the asymptotic covariance matrix for the predicted probabilities, which one can also compute using the delta method (Greene, 2003, p. 674):

$$\text{Asy. Var}[F(\boldsymbol{\beta}'\mathbf{x})] = f(\boldsymbol{\beta}'\mathbf{x})^2\mathbf{x}'\mathbf{V}\mathbf{x}. \tag{10}$$

Given the above normalizations, one can show that the asymptotic variance of the predicted probability for the reference group is $f(c)^2\sigma_c^2$, and the asymptotic variance for the dummy variable group is $f(c+d)^2(\sigma_c^2+\sigma_d^2+2\sigma_{cd})$ where σ_d^2 is the asymptotic variance for dummy variable parameter d, and σ_{cd} is the asymptotic covariance between c and d. One can further show that the asymptotic variance of the effect of the dummy variable (i.e., of the difference in predicted probability) is

$$\text{Asy. Var}[F(c+d) - F(c)] = f(c)^2\sigma_c^2 + f(c+d)^2(\sigma_c^2 + \sigma_d^2 + 2\sigma_{cd}), \tag{11}$$

which, as before, can be easily computed using Table 1 along with the estimated values of c, d, σ_c^2, σ_d^2, and σ_{cd}.

5. Concluding remarks

The need to compute marginal effects from probit and logit models rather than simply looking at raw parameter estimates is one of the few inconveniences of an otherwise extremely convenient modeling specification. We show that even this inconvenience can be minimized with an appropriate normalization of the explanatory variables. Using only the raw logit or probit estimates and the table given herein, one can compute all the desired marginal and discrete effects, along with their variances. In addition to

[4] Although this may not appear to be a dramatic simplification, we point out that without our normalizations, this calculation could potentially involve all $n \times n$ entries in the estimated covariance matrix, where n is the number of model parameter estimates. With the simplification it can be done on a hand calculator.

obviating the need for extra programming or reliance on software capabilities, this approach helps build intuition and clarifies the relationship between discrete choice parameter estimates and their associated effects.

Acknowledgements

We thank Billy Pizer, Spencer Banzhaf, and an anonymous referee for comments on a previous version of the paper, and acknowledge financial support from U.S. DOE grant DE-FG02-98ER62702.

Reference

Greene, W.H., 2003. Econometric Analysis, 5th ed. Prentice Hall, Upper Saddle River, NJ.